

Resumen

El aumento de las soluciones tecnológicas disponibles y de sus capacidades, sumado a la cristalización del fenómeno *big data* y los recientes avances de la inteligencia artificial, están concediendo a la industria financiera un mundo de nuevas posibilidades de gestión. De forma continua, las personas, los negocios y las propias máquinas están generando multitud de datos e información susceptible de generar movimientos de gran magnitud en los mercados financieros. Esto exige a los grandes gestores desarrollar sistemas que estén conectados a todas las fuentes de información, interpretando y analizando datos de forma continua, y tomando de forma autónoma decisiones inmediatas de inversión a partir de ellos. Y en todo ello, la inteligencia artificial y en particular el *machine learning*, juegan un papel fundamental. Esto está provocando que aproximadamente dos tercios de las operaciones que reciben los mercados financieros sean generadas por *bots* de forma automática. Esta importante transformación que está sufriendo el sector permite intuir que la supervivencia de las empresas del sector en los próximos años, sin duda, estará determinada por la incorporación de estas tecnologías.

Palabras clave: inteligencia artificial, industria financiera, *machine learning*.

Abstract

The increase in available technological solutions and their capabilities, coupled with the crystallization of the big data phenomenon and the recent advances in artificial intelligence, are giving the financial industry a world of new management possibilities. On a continuous basis, people, businesses and the machines themselves are generating a multitude of data and information that can generate large-scale movements in financial markets. This requires large managers to develop systems that are connected to all sources of information, interpreting and analyzing data continuously, and making autonomous immediate investment decisions from them. And in all this, artificial intelligence and in particular machine learning, play a fundamental role. As a result of this, approximately two thirds of the operations currently traded in financial markets are automatically generated by bots, underlying the fact that embracing new technologies will play a major role on companies future development and success.

Keywords: artificial intelligence, financial industry, machine learning.

JEL classification: G10.

APLICACIONES DE LA INTELIGENCIA ARTIFICIAL A LOS MERCADOS FINANCIEROS

Igor ALONSO

Accurate Quant Agencia de Valores, S.A.

Adrián CARRIO

Centro de Automática y Robótica, Universidad Politécnica de Madrid/CSIC

I. INTRODUCCIÓN

Un tuit hunde los mercados

VIERNES 23 de agosto de 2019, en plena guerra comercial entre EE.UU. y China, el gigante asiático anuncia a través de un comunicado oficial del Ministerio de Finanzas la imposición de aranceles a las importaciones estadounidenses a partir del 1 de septiembre. Esto provoca una rápida caída del 0,85 por 100 en la cotización de los futuros del índice bursátil S&P 500 en un intervalo de apenas cinco minutos.

Tres horas más tarde, el presidente de EE.UU., Donald Trump, a través de su cuenta de Twitter, anuncia una respuesta inminente en forma de nuevos aranceles, e incluso solicita a las empresas americanas que dejen de producir en China. En cuestión de once minutos, tras la publicación del tuit, el mismo índice se desploma con una caída superior al 1,82 por 100 (1).

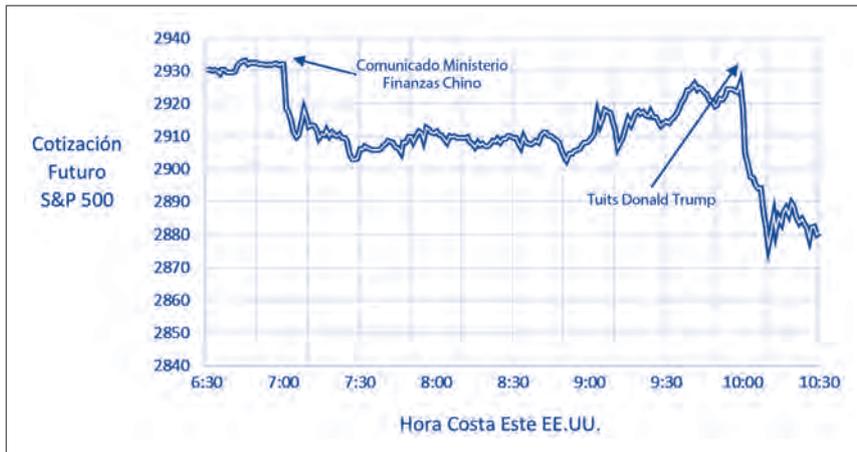
Dos formas de comunicación, dos noticias negativas, una misma consecuencia.

Vivimos actualmente en los mercados financieros una época convulsa, con valoraciones muy volátiles debido a las diversas incertidumbres que hay encima de la mesa: además de la guerra

comercial que hemos comentado, vivimos una desaceleración económica a nivel global y con la posibilidad real de un *brexit* sin acuerdo. Con ello, un simple tuit puede derivar en rápidas y fuertes correcciones del mercado.

De sobra es conocido que la información es crucial en la toma de decisiones y que el hecho de disponer de ella antes que el resto es una ventaja clara frente al mercado, luego en este sentido, si tuviésemos a nuestro alcance la posibilidad de obtener de forma instantánea el comunicado de China y el tuit de Trump, la capacidad de tener un algoritmo que pudiese interpretar de forma inmediata el contenido del mismo, en sentido positivo o negativo para el mercado, y un sistema de decisión que ejecutase operaciones de compra o venta directamente contra el mercado en función de aquella, ¿no podríamos sacar rendimiento de ello? Fijémonos en que las correcciones posteriores a las noticias, aunque se producen en cuestión de minutos, son una eternidad para un ordenador.

Estas estrategias son un ejemplo de las que, desde hace muchos años, un grupo muy selecto de grandes gestoras internacionales han implementado en sus carteras de inversión con grandes resultados, y en estas, la inteligencia artificial, como



veremos más adelante, ha sido pieza clave.

En el presente artículo veremos cómo hoy en día, gracias a las nuevas posibilidades tecnológicas, la generalización de la creación, almacenamiento y distribución de información, y a las nuevas metodologías en inteligencia artificial, las estrategias cuantitativas de gestión están ganando terreno en los mercados financieros.

II. TRANSFORMACIÓN DE LOS MERCADOS FINANCIEROS

Desde hace años estamos siendo testigos de profundos cambios en los mercados financieros y en cómo los diferentes agentes de la industria intervienen en él. Si bien en las últimas décadas presenciamos una transformación profunda en su funcionamiento, en esta década, se está produciendo un cambio notable en el *buy side* (2).

Con respecto al funcionamiento de los mercados financieros, hoy en día, prácticamente la totalidad de los activos financieros se negocian por vía electrónica. Este hecho ha supuesto mejorar radicalmente las funciones

que todo mercado debe ofrecer para su correcto funcionamiento: mejorar el acceso a los participantes, ofrecer un mayor grado de seguridad y transparencia, otorgar mayor liquidez, y finalmente mejorar la formación de precios. Este proceso de cambio fue posible gracias a la evolución tecnológica de conectividad a nivel global, con la generalización y consolidación del uso de Internet.

Por otro lado, históricamente, en las carteras de inversión ha imperado la *beta* de mercado (3) y la discrecionalidad de los gestores en sus actuaciones. Sin embargo, en los últimos años, debido a las posibilidades que ofrece la tecnología, a la cristalización del fenómeno *big data* (4) en los mercados financieros, y a las nuevas metodologías de inteligencia artificial, están ganando terreno considerablemente estrategias donde las decisiones se fundamentan únicamente en datos y probabilidades.

Los agentes financieros adoptan sus decisiones de compra y venta en función de sus necesidades y objetivos, pero fundamentalmente en base a la información de que disponen. Gran

parte de esta información, que hasta hace años era accesible de forma periódica o puntual, cada vez, además de crecer en cantidad y en calidad, se encuentra disponible en tiempo real. En este sentido, los gestores no pueden permitirse el lujo de obviar las grandes cantidades de información que la sociedad o los agentes económicos generan en cada momento, y para ello necesitan desarrollar metodologías que le posibiliten procesarla, interpretarla, analizarla y finalmente tomar decisiones en consonancia con ella. Y esto último, está viniendo de la mano de la inteligencia artificial y en particular del *machine learning*.

Esto está derivando en que, hoy en día, se considera que aproximadamente entre el 60 por 100-70 por 100 de las operaciones de los mercados financieros se realizan automáticamente por máquinas, y no por personas.

La tecnología como eje central de la evolución de los mercados financieros

No es posible entender la transformación continua de los mercados financieros sin la evolución de la tecnología. Si bien hace años la tecnología posibilitó la digitalización de los mercados financieros gracias al desarrollo y generalización de la conectividad, durante los últimos años el desarrollo tecnológico se ha centrado en el crecimiento de la capacidad computacional y de almacenamiento.

En 1965, Gordon Earle Moore, cofundador de Intel, al analizar la evolución de las capacidades computacionales entre 1959 y 1965, formuló una extrapolación, conocida como ley

de Moore, donde especulaba que la cantidad de componentes (transistores, resistencias y condensadores) integrables en un único chip se duplicaría cada doce meses (5).

Independientemente del período en el que la tecnología permite duplicar sus capacidades, la ley de Moore es una de las predicciones tecnológicas más duraderas que se han formulado, y nos da una idea de la evolución de las capacidades tecnológicas en las que nos vemos inmersos. Se considera el emblema de la era de la información, la marcha implacable del chip de ordenador que permite una revolución técnica, económica y social nunca antes experimentada por la humanidad.

Con respecto al crecimiento de la capacidad de almacenamiento de información, por ponerla en perspectiva con un ejemplo, la primera unidad de disco duro fue creada por IBM en 1956, tenía una capacidad de 5 *megabytes* (MB) y era del tamaño de una habitación. En 1980 se lanzó la primera unidad de disco duro «pequeña», con la misma capacidad de almacenamiento, pero que entraba en un ordenador personal. Diez años después, en 1990, un disco duro normal podía almacenar desde 40 a 100 MB, en función del modelo. Hoy en día, un disco duro de 10 cm, tiene una capacidad de 15 *terabytes* (TB) (15 millones de MB). Si, además, nos fijamos en el coste de almacenamiento, en los últimos treinta años, el coste de almacenamiento de gigabytes (1.000 MB) ha pasado de aproximadamente 100.000 euros a apenas unos céntimos. En resumen, de una forma exponencial, se ha evolucionado a más capacidad

de almacenamiento, en menor espacio físico y a menor coste.

Hasta que la informática cuántica sea una realidad, lo cual probablemente se produzca en un futuro próximo, y se produzca una auténtica explosión de la capacidad y velocidad computacional, lo cierto es que la actual tecnología sigue posibilitando procesar, analizar y almacenar cada vez más y más información de forma más rápida y económica.

Adicionalmente a las mayores capacidades de los ordenadores, la generalización de la informática paralela o distribuida ha supuesto del mismo modo un revulsivo considerable. La posibilidad de compartir recursos de forma remota actualmente permite incrementar las capacidades tecnológicas y de almacenamiento, lo cual también se conoce como *cloud computing*. Para el año 2020 se estima que la tercera parte de todos los datos se procesarán a través de la nube (6). Hoy en día, a través de soluciones tecnológicas *software as a service* (SaaS), *platform as a service* (PaaS) o *infrastructure as a service* (IaaS), cualquier entidad puede acceder fácilmente a estas nuevas posibilidades a un coste mucho más accesible.

III. BIG DATA EN LOS MERCADOS FINANCIEROS

La verdadera disrupción que en el mundo de la gestión se ha producido en los últimos años ha venido de la mano de la información. La toma de decisiones siempre viene condicionada por los datos que tenemos a nuestro alcance y desde hace unos años se han producido dos circunstan-

cias claves: la generación de datos ha crecido de forma exponencial y el acceso a estos se ha vuelto prácticamente instantáneo.

Aunque su capacidad de incidir en los mercados aún está vigente, la transcendencia de indicadores macroeconómicos tradicionales, tales como el cambio en nóminas no agrícolas (*non-farm payrolls*), el índice ISM (elaborado en EE.UU. por el Institute for Supply Management) manufacturero y no manufacturero, o incluso publicaciones de evolución de IPC, del PIB, comunicados de bancos centrales sobre política monetaria, han perdido fuerza frente a las nuevas fuentes de información, ya que aquellas se generan con una frecuencia baja, y en ningún caso de manera continua.

Hoy en día, la inmensa mayoría de la información que se recopila se hace de forma electrónica, dado que gran parte de nuestra actividad personal y profesional se hace *online*, y esto nos puede dar una idea de la magnitud de datos que en cada segundo se está generando a nivel global. Esto ha originado la creación de un gran negocio en torno a la generación, tratamiento y suministro del dato (7), con la aparición de multitud de entidades especializadas en múltiples campos de la información, facilitando a la industria financiera con ello el acceso a los mismos. Y esta información va desde gustos, preferencias, actividades, u opiniones de personas o colectivos, hasta información de negocios como puede ser número de visitantes a centros de negocios, transacciones de compra venta de artículos, predicciones meteorológicas sobre zonas agrícolas, situación de petroleros en diferentes partes del mundo, etcétera.

Tres características del big data: volumen, velocidad y variedad

Durante la última década, la recopilación sistemática de información, interpretación, organización y difusión han ido configurando el concepto de *big data*. Tres son las características básicas que concentra el término: volumen, velocidad y variedad.

El tamaño de los datos recopilados y registrados a través de diferentes medios es impresionante, y creciente en el tiempo, siendo el límite mínimo de referencia para ser considerado «*big*» superior cada vez. Se considera que aproximadamente el 90 por 100 de los datos almacenados actualmente han sido recopilados en los dos últimos años, y, según estimaciones, se considera que entre 2015 y 2020 se multiplicará por diez su volumen.

La segunda característica es la velocidad. Disponemos de datos prácticamente en tiempo real, bien recibidos en *streaming* (de forma continua) o en *batch mode* (por lotes). Una gran cantidad de estos datos son recopilados y distribuidos al mismo tiempo, sin apenas haber sido tratados.

La tercera característica es la variedad del formato de la información que se recibe. En función de la fuente, el formato puede ser tanto estructurado (como suele producirse en formatos de texto, CSV, SQL), en formato semiestructurado (habitual en formato HTML, JSON), o completamente desestructurado como pueden ser los procedentes de vídeos, imágenes, blogs, etc. La estructura en la recepción se vuelve clave en el tratamiento de la información, ya que añade o elimina complejidad al mismo.

La veracidad está considerada hoy en día como una característica adicional del *big data*, la cual cobra especial relevancia en los mercados financieros. Esta hace referencia a la calidad o exactitud del dato y a su valor como tal. De la calidad del dato dependerá la validez de cualquier estudio o estrategia de actuación establecida. Existen muchas bases de datos que necesitan depuraciones por poder presentar vacíos de información, duplicidades o simplemente incoherencias. Esta es una labor crítica y una de las que más recursos consume por parte de los *quants* (8).

Por otro lado, podemos clasificar la nueva generación de datos en tres apartados: datos generados por individuos, datos generados por los negocios, y finalmente los datos generados por las propias máquinas.

Big data generado por individuos

Con datos generados por individuos nos referimos a toda clase de información que personas físicas, en su día a día, en su vida privada, pueden generar. Tengamos en cuenta que una gran parte de nuestras actividades cotidianas las hacemos a través de dispositivos fijos y móviles conectados a la red, y mediante aplicaciones o *softwares* de terceros donde quedan registradas. Esta tendencia ha sido creciente en los últimos años y previsiblemente continúe con esa evolución, pese a la cada vez más restrictiva normativa de protección de datos.

Su procedencia puede ser de redes sociales, como son Twitter, LinkedIn, etc, de blogs generados por particulares, o bien por noticias o publicaciones ge-

nerales. Por otro lado, puede proceder también de entidades especializadas en recopilar información de *apps* o webs especializadas, registrando búsquedas, altas y bajas, descargas, facturación, etc. Y, por último, también puede proceder de búsquedas web y datos personales aportados voluntariamente por los usuarios, obtenidos, por ejemplo, de proveedores de gestores de correo electrónico.

Esta clase de datos suelen ser recibidos como texto, y frecuentemente de forma desestructurada, procedentes de múltiples fuentes. Este tipo de información suelen ser muy apreciada para indicar tendencias o *sentimientos de mercado* (9).

Entre los ejemplos prácticos de proveedores de esta clase de información contamos con iSENTIUM, proveedor que proporciona datos de sentimiento de mercado en tiempo real basado en mensajes de Twitter. Su información puede proporcionar a sus usuarios una herramienta de predicción de movimientos de mercado a corto plazo. Al comienzo del artículo, hablábamos de la incidencia en los mercados de un solo tuit, con mucha relevancia por proceder del presidente de EE.UU., en este caso se recogería la opinión de millones de personas que, aunque con menor incidencia de forma individual, como colectivo puede resultar un indicador adelantado extraordinario.

Otro ejemplo lo tenemos con RavenPack, proveedor que proporciona indicadores basados en noticias especializadas, novedades regulatorias, comunicados de prensa, etc. Este tipo de proveedores proporciona información que incluiría el comunicado

oficial que emitió China y al que hacíamos referencia al inicio.

Por último, también tenemos proveedores más conocidos como son Alexa o Google, los cuales nos pueden proporcionar información muy valiosa en torno a búsquedas y seguimiento web.

Big data generado por operaciones comerciales

Se trata de datos producidos o recopilados por todo tipo de entidades, públicas y privadas. Nos referimos a la información que las entidades pueden ofrecer sobre transacciones comerciales, empleo de tarjetas de crédito, afluencia a centros comerciales, variaciones de inventarios, cadenas de suministro de mercancías, etc. O al conjunto de grandes agrupaciones de datos generados por entidades públicas. Este tipo de información se presenta de una forma más estructurada que la proporcionada por los individuos y sus datos son considerados como buenos indicadores adelantados de métricas corporativas.

En este caso, tenemos como ejemplo de datos de organismos públicos como WTO (World Trade Organization) o IMF (International Monetary Fund), que suministran datos utilizados para estrategias cuantitativas a más largo plazo y de baja frecuencia. Otro ejemplo de información de operaciones comerciales la tenemos con EagleAlpha, que suministra gran variedad de información, como por ejemplo información sobre transacciones *online* de empresas integrantes del índice bursátil S&P 500.

Big data generado por sensores

Por último, tenemos los datos generados por sensores: da-

tos recopilados de forma automática a través de sensores integrados en diversos dispositivos. Estos datos son generados fundamentalmente de forma no estructurada y su tamaño es con frecuencia sustancialmente más grande que los flujos de datos generados por humanos o procesos. El ejemplo más conocido es la imagen vía satélite que puede usarse para monitorear actividades económicas, como la evolución de reservas de materias primas, de construcciones, envíos de estas o de materiales, producción de artículos, etc.). Otro ejemplo son los datos procedentes de la geolocalización que pueden ser empleados por ejemplo para rastrear el tráfico peatonal en tiendas minoristas, en aparcamientos, o de barcos de mercancías en puertos, etc. Más ejemplos tenemos con la información procedente de cámaras fijas en lugares de interés, sensores climáticos y de contaminación, etc. Con total probabilidad, es el futuro concepto de Internet de las cosas (IoT), que consiste en la práctica de incorporar microprocesadores y tecnología de red en todos los dispositivos electrónicos personales y comerciales.

Entidades como Orbital Insights, Descartes Labs, o Rezatec son entidades que pueden ofrecer información muy valiosa para establecer estrategias sobre materias primas como petróleo o bienes agrícolas.

Aspectos claves del big data para los mercados financieros

Uno de los aspectos más relevantes a tener en cuenta es la relación que puede establecerse entre la información y los mercados financieros a los que

hace referencia. La mayoría de la información que hoy en día se encuentra accesible al gestor está directa o indirectamente relacionada con acciones o sectores empresariales o con materias primas, siendo menor, por ejemplo, la información generada en torno a tipos de interés y divisas, la cual se deriva más de información macroeconómica.

Otro aspecto importante es la consideración del potencial de la información, traducido en el potencial *alfa* (10) que puede aportarnos y su coste. En primer lugar, debemos considerar el coste de disponer de datos: no es igual de costoso adquirir bases de datos que nos puedan aportar sentimiento de mercado (información más pública) que otra que nos aporte información sobre transacciones con tarjeta de crédito (con un carácter mucho más privado). A esto hay que añadir, que el coste únicamente se justifica si la rentabilidad que vamos a obtener de ella es suficiente, y eso es difícil de conocer *a priori* (11). Lo cierto es que, aunque la intuición puede indicarnos que cierta información puede ser muy útil para identificar grandes oportunidades de inversión, es realmente complicado encontrar bases de datos que nos permitan por sí solas diseñar estrategias de inversión que aporten una ratio de Sharpe (12) muy elevado a nuestras carteras.

Otro factor a tener en cuenta es lo conocida o accesible que resulta la base de datos; en este sentido, existen ciertas bases de datos que hacen referencia a magnitudes fundamentales de empresas como son beneficios o ventas, donde el potencial es mucho menor, ya que esa información es más conocida y puede

estar disponible por otras fuentes. En cambio, otras bases de datos otorgan una información mucho más exclusiva, por lo que su potencial es, *a priori*, mucho más atractivo.

Otro elemento a considerar es el estado de procesamiento en el que se encuentra la información. Podemos disponer de bases de datos con información presentadas en formato informe o resumen, o bases de datos con información totalmente desestructurada. En este último caso, el tratamiento de esta información es sustancialmente más costoso que en el primero, y eso resulta un factor determinante en la rentabilidad de la adquisición.

Probablemente, el factor más importante es la calidad de los datos. En primer lugar, debemos considerar que el tamaño de la información es trascendental; cuanto más histórico dispongamos, más robustas serán las conclusiones que podamos obtener, y esto es algo de lo que no se dispone en todas las bases de datos (13). En segundo lugar, debemos tener en cuenta la forma en la que se ha obtenido (metodología) y si existen gaps de información en el historial (en este caso es trascendental conocer el porqué, si han sido aleatorios o existe algún patrón, y cómo se han solucionado).

Por último, deberíamos considerar también la disponibilidad de ella, no es lo mismo disponer de la información en tiempo real, de forma diaria, semanal, mensual, etc. Del mismo modo, no es lo mismo si existe latencia o si se producen interrupciones, ni la forma de conexión al proveedor, ya que pueden existir grandes diferencias de robustez y seguridad entre conexiones.

Recordemos la importancia que en nuestro ejemplo inicial tendría disponer de la información con unos minutos de retraso.

En definitiva, cada gestor dará más importancia a unos aspectos que a otros en función de su estilo de gestión; en unos casos, la inmediatez será clave, por ejemplo, para los gestores de alta frecuencia, pero en cambio gestores con un enfoque más tendencial ponderarán más la singularidad y exclusividad del dato.

IV. MACHINE LEARNING – INTELIGENCIA ARTIFICIAL

Una mirada histórica

En los últimos años ha habido una evolución significativa en el campo del aprendizaje automático o *machine learning*, un conjunto de técnicas que, combinando algoritmos matemáticos con modelos estadísticos, permiten a un sistema informático resolver tareas sin instrucciones explícitas, empleando fundamentalmente datos. El *machine learning* combina, por tanto, las disciplinas estadística e informática y se engloba dentro de un conjunto más amplio de técnicas, en lo que actualmente se conoce como inteligencia artificial.

La inteligencia artificial no es un área de conocimiento tan joven como pueda parecer. Nació como disciplina académica en el año 1956 durante una célebre conferencia en Dartmouth, Estados Unidos, a la que asistieron los principales expertos de aquel momento, en la que se definieron sus objetivos y donde se hicieron públicos sus primeros éxitos.

Durante los primeros años de la inteligencia artificial se resol-

vieron problemas algebraicos, se demostraron teoremas geométricos y se implementaron los primeros sistemas que podían hablar en inglés. Estos avances generaron un gran optimismo (se creía en la posibilidad de replicar el cerebro humano en pocos años) y todo ello propició la inyección de grandes inversiones para acelerar su desarrollo. Sin embargo, a partir del año 74 y hasta el año 80, las grandes expectativas creadas se vieron superadas por la realidad, y la lentitud con la que avanzaban los desarrollos generó gran decepción y recortes en las inversiones.

Posteriormente, entre los años 1980 y 1987 tuvo lugar un segundo *boom* con la aparición de los llamados «sistemas expertos» o sistemas basados en reglas que permitieron resolver multitud de problemas empleando reglas lógicas y el conocimiento de personas expertas. La implantación de estos sistemas generó grandes ahorros en las empresas que los adoptaron y también provocó una gran aceleración en las empresas de desarrollo de *software* y de fabricación de *hardware* informático. Además, los avances en el desarrollo de las redes neuronales artificiales, cuya utilidad había sido discutida hasta ese momento, impulsaron su uso de manera definitiva. Todo esto generó de nuevo gran optimismo y movilizó grandes inversiones que apostaban por los avances de la inteligencia artificial.

A finales de los años ochenta y a principios de los noventa, la aparición del ordenador personal provocó el hundimiento de las empresas de *hardware* especializado, un sector de enorme peso económico. Esto, sumado a los problemas para mantener los sis-

temas expertos y a las enormes expectativas que se habían generado una vez más, provocó de nuevo decepción entre los inversores y grandes recortes. Sin embargo, debido a los avances obtenidos desde mediados de los años noventa en multitud de áreas, minería de datos, robótica, reconocimiento del habla, diagnóstico de enfermedades, etc., vivimos ya una tercera revolución de la inteligencia artificial y el optimismo ha surgido de nuevo con la resolución de algunos de los retos que se habían planteado en su nacimiento, como vencer al campeón del mundo de ajedrez o manejar de forma segura un coche sin conductor.

Desde el año 2011 con la disponibilidad de grandes volúmenes de datos, de ordenadores más potentes y de técnicas de *machine learning* más avanzadas, la inteligencia artificial se viene aplicando en un número cada vez mayor de problemas. En concreto, durante este último período se ha producido el nacimiento del *deep learning*, que se refiere al aprendizaje automático con estructura jerárquica, como el que se produce cuando se aprende a reconocer un objeto en una imagen empleando varios niveles de abstracción: reconociendo partes de dicho objeto, su estructura y/o su entorno. El avance del *deep learning* estos últimos años ha permitido grandes avances no solo en el ámbito financiero, sino también en el procesamiento de imagen, vídeo, el análisis de texto y el reconocimiento del habla.

Machine learning

Centrándonos en los algoritmos de aprendizaje automático o *machine learning*, se dice que un algoritmo «aprende» a resolver

una tarea (por ejemplo operar un determinado activo financiero) cuando existe una forma de medir el desempeño de dicha tarea (por ejemplo, midiendo la rentabilidad diaria obtenida) y dicho algoritmo mejora su desempeño conforme adquiere experiencia (por ejemplo, operando el activo financiero durante un período de tiempo, es posible inferir automáticamente una estrategia que ofrezca resultados favorables).

Un mismo problema (operar un activo financiero) se puede plantear mediante diferentes tareas (prediciendo precios, detectando patrones, etc.) y es posible utilizar diferentes métricas para medir su desempeño (rentabilidad, pérdida máxima acumulada, ratio de Sharpe, etc.). Además, la experiencia puede ser adquirida de diferentes formas, las cuales se clasifican en tres grupos principales que dan nombre a los tres tipos de aprendizaje automático más comunes: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo.

Aprendizaje supervisado

El aprendizaje supervisado o *supervised learning* se basa en tratar de imitar la inferencia que hace una persona experta para resolver una determinada tarea. Por ejemplo, si la tarea a resolver consiste en clasificar una serie temporal como tendencial o en rango, el algoritmo debe procesar un conjunto de series junto con la información sobre cuáles son tendenciales y cuáles en rango. Esto último recibe el nombre de etiqueta o *label* y es el conocimiento que aporta el experto. Una vez entrenado con estos datos, el algoritmo será capaz de clasificar una nueva serie que no ha visto antes. Otra de las ta-

reas más habituales que resuelven este tipo de algoritmos junto con la clasificación es la regresión, en la cual el algoritmo no busca predecir una clase (por ejemplo, tendencial o en rango), sino el valor numérico de una o varias variables. Por ejemplo, si la tarea es predecir el precio de cierre de hoy a partir de los precios de cierre de los últimos días, el algoritmo debería entrenarse usando varios bloques de mil precios de cierre junto con el precio de cierre del día posterior a cada bloque. Este último precio de cierre sería la etiqueta de cada uno de los bloques. Como ya se mencionó anteriormente, en la actualidad se utilizan grandes volúmenes de datos, incluyendo no solo precios o indicadores técnicos, sino también información como los sentimientos de mercado, que pueden extraerse automáticamente de las noticias empleando técnicas de procesamiento del lenguaje natural y lo que se conoce como análisis del sentimiento o minería de opinión.

Para aplicar con éxito todas estas técnicas no solo habrá que elegir correctamente el mejor algoritmo (redes neuronales artificiales, árboles de decisión, SVMs –máquinas de soporte vectorial, *support vector machines*–, etc.) junto con sus parámetros, sino también realizar un estudio que garantice que no existe sesgo en el aprendizaje y que el algoritmo es capaz de generalizar correctamente cuando se le presentan datos que no ha visto anteriormente.

Aprendizaje no supervisado

El aprendizaje no supervisado o *unsupervised learning* recibe este nombre ya que no requiere la supervisión de un experto que proporcione etiquetas, puesto que sus algoritmos son capaces

de encontrar automáticamente patrones en los datos. Algunas de las tareas más habituales que resuelven este tipo de algoritmos son la detección de anomalías y el *clustering*, en el cual se busca agrupar los datos en clústeres o grupos, de manera que los datos clasificados en un mismo clúster compartirán determinadas características y se diferenciarán de alguna manera de los datos presentes en el resto de clústeres.

Un ejemplo de la aplicación del aprendizaje no supervisado en los mercados financieros podría ser la búsqueda de patrones dentro de una serie de datos. Si tomamos la serie y la dividimos en bloques de longitud fija, podríamos utilizar un algoritmo de *clustering* para agrupar los bloques en varios grupos, de manera que una vez completado el aprendizaje cada clúster podría entenderse como un patrón y los clústeres con mayor número de bloques contendrían los patrones más frecuentes.

La clave en este tipo de algoritmos es determinar cómo se mide la distancia entre dos muestras cualesquiera, de manera que se codifique correctamente la similitud entre ellas. Como referencia, algunos de los algoritmos más comunes en el aprendizaje no supervisado son el algoritmo de *k-means*, los *autoencoders*, los mapas autoorganizados (*SOM*, *self organizing maps*) o el análisis de componentes principales (*PCA*, *principal component analysis*).

Aprendizaje por refuerzo

Por último, el aprendizaje por refuerzo o *reinforcement learning* permite aprender a resolver una tarea mediante un sistema de castigos y recompensas. Durante

el aprendizaje, un agente interactúa con un entorno (real o simulado) observando una serie de variables conocidas como «estado» y ejerciendo acciones sobre dicho entorno. Esas acciones provocan una serie de castigos y recompensas prefijados por el usuario en lo que se conoce como función de recompensa o *reward function*. Inicialmente, el agente toma decisiones de manera prácticamente aleatoria, pero conforme recibe la realimentación de sus acciones, es capaz de discernir cuáles pueden llevarle a maximizar su recompensa y en qué momento es mejor aplicarlas. El algoritmo debe ponderar la exploración de nuevas estrategias, que le permitirán mejorar su desempeño, con la explotación de aquellas que ya ha descubierto. El objetivo final es que el agente genere por sí mismo un sistema de reglas según el cual sepa en cada momento, dado un estado, tomar la acción que le permita obtener las máximas recompensas posibles.

Un ejemplo de la aplicación de este tipo de algoritmos a los mercados financieros podría ser la gestión de una cartera, en la cual un gestor o *trader* debe decidir si compra o vende un determinado activo financiero en función de los recursos disponibles para invertir y del estado del mercado. El agente (que en este caso equivaldría al *trader*) podría entrenarse en un entorno de simulación antes de operar en el mercado real. En dicho entorno de simulación se simularía la fluctuación del valor de los activos (con datos reales o simulados) y de la cartera del agente, lo cual constituiría el estado (al cual se podrían añadir otras variables de interés como por ejemplo indicadores técnicos de los activos), en base al cual el agente toma sus decisiones. Las acciones serían el lanzamiento

de órdenes de compra o venta de activos, las cuales tras ejecutarse generarían beneficios y pérdidas, que se traducirían en recompensas y castigos en base a la función de recompensa fijada de antemano, que podría ir encaminada, por ejemplo, a obtener el máximo beneficio con la mínima exposición. Como se ha explicado anteriormente, el agente comenzaría operando de manera aleatoria, pero después de determinado número de operaciones, el agente sería ya capaz de tomar decisiones que tendiesen a generar recompensas más altas cada vez en base a la experiencia adquirida.

Una de las grandes fortalezas del aprendizaje por refuerzo es que permite ponderar el beneficio inmediato frente el beneficio a largo plazo. El objetivo perseguido por el agente no tiene por qué ser necesariamente obtener el máximo beneficio de manera inmediata tras una acción, sino el beneficio en el largo plazo, aunque ello conlleve un «castigo» inmediato. Ejemplos de esto serían el incurrir en costes de operación para abrir una posición que se espera cerrar con beneficios o el cerrar una posición con el objetivo de minimizar una pérdida.

Entre los algoritmos de aprendizaje por refuerzo más utilizados destacan aquellos que aprenden *off-policy*, es decir, ignorando las reglas aprendidas hasta ese momento, como por ejemplo *Q-Learning* y aquellos que sí la tienen en cuenta como *state-action-reward-state-action (SARSA)*, *non-adjacent form (NAF)* y *proximal policy optimization (PPO)*, conocidos como algoritmos *on-policy*.

Los avances recientes en el campo de la inteligencia artificial

y en particular en el *machine learning* están poniendo herramientas muy potentes a disposición de los profesionales de las finanzas. La rápida y correcta adopción de estas herramientas, sin duda, contribuirá a la toma de decisiones financieras más informadas, controlando su riesgo de manera mucho más eficaz.

La máquina supera al hombre

Con la inteligencia artificial, en definitiva, lo que se pretende es replicar el proceso cognitivo humano, donde se recibe información, se procesa y analiza, se adoptan decisiones, y se sacan conclusiones en función de los resultados. En este sentido, el objetivo es claro y alcanzable, en la medida que seamos capaces de descifrar todos los condicionantes que rodean al ser humano en su toma de decisiones, tanto sean considerados racionales como emocionales. Un *bot* (14) debería ser capaz de replicar las decisiones del ser humano con mayor o menor dificultad. Incluso, si el proceso de réplica cognitiva permite discernir ciertos condicionantes que podemos considerar perjudiciales en la toma de decisiones, por ejemplo, las más puramente emocionales o sesgadas, podría mejorar al ser humano, más allá de las ventajas evidentes que una máquina puede aportar por el simple hecho de poder tratar muchísima más información en un tiempo menor, y contemplando muchísima más casuística.

Procesos de réplica humana «mejoradas» se han producido en los últimos años en diversos ámbitos, como por ejemplo el caso de la supercomputadora desarrollada por IBM Deep Blue que consiguió batir al campeón del mundo de ajedrez Garri

Kaspárov en 1997. Este ordenador, a través de algoritmos de inteligencia artificial, gracias a su gran capacidad de cómputo, era capaz de profundizar en nodos, es decir, en búsquedas de soluciones tras cada jugada. Era capaz de calcular 200 millones de posiciones por segundo, con un objetivo claro, maximizar las ganancias y minimizar las pérdidas de cada jugada (15). Debemos considerar que se trata de un juego de estrategia basado en piezas y casillas finitas, y por tanto acotables (16). Esto, ya de por sí, concede una ventaja al ordenador, en la medida de disponer de un universo finito de posibilidades parametrizables.

Otro caso reseñable de la competición del ser humano contra la máquina es el caso del juego Go, juego de estrategia de origen chino con una antigüedad de más de 2.500 años, y que en la antigüedad fue considerado como una de las cuatro artes esenciales de la antigua China. Se trata de un juego complejo de dos jugadores, donde la estrategia es clave. Sobre este juego, Google Deepmind, división especializada de Google, desarrolló un *software* denominado por AlphaGo (17) capaz de ganar al campeón del mundo de Go. A través de una metodología de inteligencia artificial basada en redes neuronales, desarrollaron un *software* basado en aprendizaje supervisado por jugadores expertos, reforzando el aprendizaje jugando contra sí misma.

Por último, y por poner un ejemplo de superación de la máquina al hombre en el campo del lenguaje y del conocimiento, tenemos el caso de Watson de IBM donde batió a los mejores concursantes de Jeopardy, un concurso de preguntas

y respuestas de la televisión de Estados Unidos de conocimiento general. En sí, el *software* desarrollado por IBM tenía acceso a prácticamente ilimitadas fuentes de información, bien a través de bases de datos, enciclopedias y acceso a Internet, pero realmente, el hecho diferencial es que fuese capaz de acceder, seleccionar y procesar la información más adecuada en cada momento, a través de interpretar preguntas en un lenguaje coloquial, es decir, siendo capaz de interpretar dobles sentidos, ironías y juegos de palabras, así como ser capaz de detectar matices del lenguaje ambiguo cuando el significado depende del contexto.

Estamos ante tres casos en los que un *software* desarrollado con inteligencia artificial es capaz de superar al hombre. Se trata de eventos donde el objetivo es replicar el proceso cognitivo del ser humano, pero mejorando los resultados, al disponer de más información de forma más rápida.

V. TRADING CON INTELIGENCIA ARTIFICIAL

La aplicación de la inteligencia artificial no es algo nuevo en la industria financiera, aunque en los últimos años está adquiriendo un protagonismo especial. Como comentábamos en la introducción del presente artículo, de hecho, existen entidades que desde hace décadas vienen implementando estrategias que de alguna manera incluyen inteligencia artificial en su gestión.

Trading algorítmico

Los cimientos de la inteligencia artificial en los mercados financieros vienen de la mano del *trading* (18) algorítmico.

Se trata de estrategias que, gracias a potentes ordenadores e información histórica, emplean complejos algoritmos matemáticos para tomar sistemáticamente decisiones de inversión. Estos algoritmos pueden basarse en múltiples variables técnicas de precios, bien en ratios fundamentales, ratios macroeconómicas, o cualquier otra variable cuantificable de la que se disponga y con suficiente información histórica para contrastarlo empíricamente.

La gran explosión de este tipo de estrategias de *trading* se produjo en los años ochenta, cuando las principales bolsas empezaron a posibilitar el enrutamiento automático de órdenes y las grandes instituciones de inversión empezaron a invertir en ordenadores con grandes capacidades computacionales. Aunque por aquel entonces este tipo de estrategias estaban limitadas a unas cuantas entidades, con el tiempo se han ido generalizando hasta alcanzar porcentajes muy elevados.

Hay casos especialmente exitosos de empresas que han triunfado empleando estrategias de *trading* algorítmico, como es el caso de Renaissance Technologies LLC, *hedge fund* –fondos de cobertura o fondos de inversión libre– fundado en 1982 por James Simons, un matemático galardonado y exdescifrador de códigos de la Guerra Fría. Gracias a su fondo Medallion, que se gestiona en base a modelos matemáticos estadísticos desarrollado por reconocidos físicos y matemáticos sin perfil económico, ha sido considerado como el gurú de la gestión cuantitativa y uno de los mejores *traders* de la historia.

Otros casos similares son, por ejemplo, D. E. Shaw & Co., L.P. entidad fundada en 1988 por

David E. Shaw conocida por desarrollar sofisticados modelos matemáticos y avanzados programas informáticos especializados en explotar anomalías en los precios de los mercados. O por ejemplo Two Sigma Investments, fundada en el año 2001 por John Overdeck, David Siegel y Mark Pickard.

En todos estos casos, sus estrategias están basadas en métodos cuantitativos de gestión y, según ellos mismos reconocen, en inteligencia artificial, por lo que pueden considerarse que son una evidencia de la aplicabilidad con éxito de la gestión a través de algoritmos.

Dentro de este tipo de estrategias, detengámonos en una con un enfoque más inmediato: *high frequency trading*.

High frequency trading

El *high frequency trading* consiste en la negociación de activos financieros mediante la generación de miles y miles de órdenes de compra y venta enviadas al mercado en fracciones de segundos. Estas órdenes tienen una vida media muy corta, ya que buscan pequeñas oportunidades de beneficio en cada operación. Esta forma de operar no se basa en predecir el valor futuro de los activos financieros, sino aprovechar pequeñas ineficiencias del mercado u oportunidades que ofrece la formación de precios en determinados momentos. Esto requiere el empleo de *bots* o estrategias automatizadas que permitan detectar oportunidades de negocio instantáneas, mediante ordenadores con extraordinarias capacidades computacionales necesariamente ubicados muy cerca de los centros de negociación, en aras de tener una

conexión instantánea sin apenas latencia. Estas estrategias automatizadas incluyen complejos algoritmos que analizan el precio de los activos y, sobre todo, el libro de órdenes existente en cada momento.

Estas estrategias necesitan analizar mucha información en muy breve espacio de tiempo, información que, gracias a las nuevas fuentes de información, cada vez son mayores y más diversas y precisan de nuevas tecnologías de interpretación y análisis, como las que ofrece la inteligencia artificial.

Este tipo de estrategias están mal vistas por la opinión pública: se considera que manipulan los mercados, que tienen ventajas notorias frente al pequeño inversor (de conexión y computacional), que su finalidad es puramente especulativa, y finalmente porque son capaces de mover los mercados de forma muy peligrosa.

Prueba de esto último son los famosos *flash crash*, periodos cortos de tiempo donde el precio de un activo puede caer bruscamente en cuestión de minutos, para luego recuperarse inmediatamente. Uno de los más conocidos es el que se produjo en el Dow Jones el 6 de mayo de 2010 donde sufrió una caída del 9 por 100 de su cotización para prácticamente recuperar su valor inicial en apenas quince minutos (19).

Sin embargo, frente a esta corriente contraria a este tipo de *trading* están los que ponen encima de la mesa las ventajas que aportan: en gran medida son creadores de mercado que dan liquidez en ciertos momentos de mercado y para activos ilíquidos.

Según datos de la consultora TABB Group el *trading* de alta frecuencia supone el 41,5 por 100 de todas las operaciones que se efectúan en el mercado. Del resto de operaciones, un 15 por 100 serían efectuadas en base a modelos matemáticos y estadísticos, un 16,5 por 100 por *hedge funds*, y el 26,5 por 100 en base a criterios convencionales (fundamentales).

Inteligencia artificial: dos vertientes

Las aplicaciones de la inteligencia artificial en el *trading* presentan dos vertientes: en la interpretación del dato y en la toma de decisiones.

Volviendo al caso real que hemos visto al inicio del artículo, quedó claro que el hecho de disponer de información y ser capaz de interpretarla antes que nadie puede darnos una ventaja competitiva muy lucrativa frente al mercado. Sin embargo, la inteligencia artificial juega un papel trascendental en dos ámbitos: primero en la interpretación del comunicado y del tuit, y posteriormente en la toma de decisión de compra y venta posterior.

Si bien en la interpretación del comunicado y del tuit podemos afirmar con rotundidad que en ambos casos son mensajes negativos para lo que el mercado estaba esperando, y en este sentido puede considerarse factible desarrollar un algoritmo que sea capaz de detectarlo (20), en el caso de la reacción del mercado, el algoritmo estaría haciendo una predicción futura sobre cómo los mercados reaccionarán, donde la incertidumbre es mayor.

Aunque en el ejemplo anterior pudiese parecer obvio que ante una noticia así los merca-

dos se iban a desplomar, esto no siempre ocurre: a veces, el mercado reacciona de forma inesperada (21). Por este motivo, a pesar de las grandes facilidades que la tecnología, el *big data* y las nuevas metodologías de *machine learning* ofrecen a la industria de gestión, debemos reflexionar si realmente tenemos hoy en día ante nosotros las herramientas para encontrar el santo grial o tenemos más medios para concluir, con más evidencias, que los mercados son aleatorios e impredecibles.

La máquina frente al gestor

Sobre la incógnita de si los mercados son predecibles o no, la teoría del paseo aleatorio, contemplado por Burton G. Malkiel en su obra *Un paseo aleatorio por Wall Street*, nos viene a cuestionar la validez de cualquier tipo de información que podamos obtener del propio valor, siendo un sinsentido enfocar nuestros recursos y esfuerzos en buscar información que realmente aporte valor a nuestro estilo de gestión. Su autor, reconocido economista, escritor y profesor, plantea que los mercados son eficientes, incluyendo en su precio cualquier tipo de información existente, por lo que cualquier estrategia que no sea *buy and hold* (22) de una cartera que replique los mercados de una forma diversificada, no justificaría sus costes, ya que ninguna otra estrategia activa (técnica, fundamental o discrecional) consigue batir de una forma estable a los mercados.

Evidentemente, si consideramos que los mercados son aleatorios o simplemente caóticos, cualquier esfuerzo en intentar parametrizar o reconocer ciertos patrones de comportamiento, sería una pérdida de tiempo,

independientemente de que utilicemos las mejores bases de datos, dispongamos de ellas en tiempo real o dispongamos de las últimas y más desarrolladas metodologías de *machine learning*. Sin embargo, si consideramos que existen ciertos patrones que se repiten a lo largo del tiempo, o que existen relaciones estadísticamente representativas entre diferentes variables, podríamos desarrollar algoritmos que concediesen al gestor una ventaja frente al mercado, y podría sacar provecho de ella.

Los mercados financieros consisten en la confluencia de innumerables participantes, con intereses, objetivos, necesidades, limitaciones, conocimientos, información, experiencia y un largo etcétera, muy heterogéneos, lo que provoca que las transacciones y sus valoraciones de los activos puedan ser muy diferentes y cambiantes a lo largo del tiempo. Independientemente de la discusión que podamos tener sobre la diferencia entre precio y valor (23), el precio de las cosas es algo objetivo y el valor subjetivo, sin embargo, el valor (subjetivo) de las cosas es lo que determina el precio (objetivo) que los participantes están dispuestos a ofrecer o demandar por los activos.

Siguiendo esta línea argumental, lo cierto es que entonces no puede negarse que los mercados tienen cierto grado de aleatoriedad. Pero tampoco puede negarse que se producen ciertos patrones detectables en la evolución de los precios de los mercados: los precios se crean por la confluencia de oferta y demanda según valoraciones, como decíamos, subjetivas, donde la psicología y las emociones humanas juegan un papel clave en su formación, y el ser humano, en su

comportamiento, presenta constantemente patrones repetitivos. Ni tampoco puede negarse que existan relaciones entre variables en las que se pueda acreditar una conexión causa-efecto (como en el ejemplo inicial del comunicado y el tuit), o simplemente que existan ciertas ineficiencias en la formación de precios, aunque mínimas, arbitrables (24).

Es decir, a diferencia de lo que considerábamos con el ajedrez y el Go, donde todas las posibilidades eran teóricamente calculables, en los mercados financieros no, no hay casillas ni fichas finitas, ni como con Jeopardy, donde la interpretación del lenguaje tiene un menor grado de relatividad, pero ello no implica que no existan oportunidades.

En definitiva, siempre pueden producirse ciertas circunstancias no parametrizables ¿podría una máquina anticiparse o prever las posibles consecuencias de una guerra nuclear que nunca ha ocurrido? Sin embargo, hay evidencias de que se producen en los mercados ciertas circunstancias o patrones que se repiten de forma repetitiva a lo largo del tiempo de una forma estadísticamente representativa, o ciertas informaciones que condicionan directamente el valor de los activos, concediendo cierta ventaja si se conocen con anterioridad (sin referirnos a *inside trading* [25]), o existen, aunque cada vez menos debido al desarrollo de los mercados, ciertas oportunidades o arbitrajes escondidos entre valoraciones de mercado.

VI. CONCLUSIONES

Tal como comenzábamos el artículo, un simple mensaje publicado en una red social es

capaz de arrastrar los mercados casi un 2 por 100 es cuestión de minutos, por lo que un inversor institucional no puede permitirse el lujo de no disponer de esta información de forma inmediata. Esta necesidad es la que está provocando en los mercados financieros el auge de técnicas de gestión cuantitativas basadas principalmente en tecnología, *big data* e inteligencia artificial.

A la espera de una nueva era tecnológica con la llegada de la informática cuántica, en los últimos años estamos siendo testigos de incrementos exponenciales de las capacidades computacionales y de almacenamiento, y de nuevas soluciones informáticas virtuales distribuidas, que han incrementado las posibilidades de la industria financiera, alcanzando cotas inimaginables hace años.

Vivimos en una época donde, de forma continua, las personas, los negocios y las propias máquinas están continuamente generando datos e información. La cristalización del fenómeno *big data* está facilitando la aparición de múltiples proveedores capaces de ofrecer esta información de forma legible para la industria en tiempo real y con ello posibilitando a los gestores nuevas herramientas de decisión.

La inteligencia artificial se ha vuelto clave en este punto, puesto que estas inmensas cantidades de información han de ser tratadas y posteriormente analizadas de forma automática, por lo que el desarrollo de sistemas basados en *machine learning* capaces de efectuar esta tarea al instante, se ha vuelto imprescindible. Del mismo modo, se hacen imprescindibles la elaboración de sistemas automáticos de *trading*

que de forma inmediata adopten decisiones de inversión, y aprovechen las oportunidades que el *big data* ofrece, y aquí, de nuevo, el *machine learning* juega un papel clave.

Con ello, cada vez es más habitual ver estadísticas con porcentajes muy elevados de operaciones efectuadas por máquinas en lugar de personas, sistemas robotizados que participan en los mercados de forma automática, cruzando órdenes contra otros sistemas. Sin embargo, pese a lo que podamos pensar, no estamos cerca de lo que se denomina singularidad tecnológica (26).

Constantemente hay evidencias de que el ser humano está detrás de los grandes movimientos de los mercados. Warren Buffet, considerado por muchos como el mejor inversor de la historia, y no habiendo utilizado inteligencia artificial en sus decisiones, tiene una frase célebre sobre los mercados de valores, definiéndolos como «maniaco-depresivos». Esto lo justifica con que puntualmente los mercados reaccionan bruscamente sin ninguna base lógica, tan solo dejándose llevar por la euforia y el pánico, emociones intrínsecamente humanas. De ahí que sigan creándose y explotándose burbujas como la de las criptomonedas del pasado ejercicio, o de las «sorprendentes» reacciones del mercado ante acontecimientos tales como el referéndum del *brexit*, que marcaron pérdidas históricas iniciales para recuperarlas en mes y medio, o el desplome inicial de los mercados ante los primeros sondeos que daban la victoria a Donald Trump, para terminar la sesión con ganancias cuando se confirmó la noticia.

También debemos tener en cuenta que la estrategia más extendida a nivel global es la que señalábamos al comienzo del artículo, de *buy and hold*, es decir estrategias de inversión cuya finalidad es invertir y mantener la posición en los mercados, estrategias no muy activas en su operativa, más allá de rebalances o coberturas.

En definitiva, podemos concluir que la inteligencia artificial se ha hecho un hueco en los mercados financieros como una herramienta imprescindible en la industria de inversión, por lo que su uso se generalizará con seguridad en los próximos años. Sin embargo, la inteligencia artificial difícilmente sustituirá al gestor humano totalmente, sino que, salvo casos puntuales de entidades muy especializadas, en general lo complementará en su toma de decisiones o se empleará como instrumento para crear estrategias colaterales a la inversión principal.

NOTAS

(1) Por dar una proporcionalidad a este dato, el índice VIX, que representa la volatilidad o movimiento esperado del índice S&P 500 en un año, está en niveles en torno a 18, lo que quiere decir que, en once minutos, el índice tuvo el 10 por 100 del movimiento esperado por el mercado para un año entero.

(2) *Buy side* hace referencia a las entidades financieras que adquieren servicios o productos financieros, tales como fondos de inversión o pensiones, compañías de seguros, *hedge funds*, etcétera.

(3) La *beta* de mercado hace referencia a que, de una forma mayoritaria, las estrategias de inversión de la industria han seguido la evolución de los mercados de una forma cuasipasiva, por lo que sus rentabilidades y riesgos han venido determinados por los mercados en sí y no tanto por las decisiones de los gestores.

(4) El término *big data* hace referencia a grandes y diferentes conjuntos de información que aumentan a tasas cada vez mayores. Incluye tres aspectos claves: el volumen de

información, la velocidad a la que se crea y recopila, y la variedad o alcance sobre los datos que incluye.

(5) Este pronóstico se hizo realidad en la década siguiente, y en 1975, con más información, revisó la estimación del período de duplicación a veinticuatro meses. Por aquel entonces, los componentes también permitirían duplicar la velocidad del chip, ya que la mayor cantidad de componentes podía realizar operaciones más potentes y los circuitos más pequeños permitían velocidades de reloj más rápidas.

(6) Ya son muchos los proveedores tecnológicos que emplean esta metodología (en el caso de Google, por ejemplo, emplea aproximadamente 1.000 ordenadores de forma paralela para dar respuesta a sus servicios).

(7) Se estima que en la actualidad el mercado de *big data* representa aproximadamente 130.000 millones de dólares americanos, cifra que se espera que alcance los 200 en el año 2020. La industria financiera es uno de los grandes participantes, representando aproximadamente el 15 por 100 del total, según estimaciones de JP Morgan.

(8) Con el término *quant* se hace mención a los profesionales expertos en el análisis y gestión de información cuantitativa. En la industria financiera desempeñan un papel crítico ya que reúnen amplios conocimientos estadísticos y matemáticos, y por otro lado conocimientos en diferentes lenguajes de programación.

(9) Como sentimiento de mercado se considera la opinión generalizada de los inversores sobre un activo o mercado. El sentimiento de mercado condiciona en gran medida la evolución de los precios de los activos y suele tener un componente emocional muy importante.

(10) *Alfa* es considerada la rentabilidad que puede aportarnos como inversores independientemente de la evolución de los mercados (esta última viene reflejada con la *beta* a la que previamente hacíamos alusión).

(11) Para estos casos suelen realizarse muestreos de información con los que se elaboran estudios estadísticos denominados *backtestings*, pero no siempre resultan concluyentes.

(12) Ratio financiera que mide la rentabilidad de una cartera por encima de la rentabilidad del *active* libre de riesgo, con respecto al riesgo (volatilidad) que se ha asumido para lograrla. Cuanto mayor sea la ratio, más atractiva es la inversión.

(13) El tamaño de la información depende en gran medida de la naturaleza del dato y el proveedor. Como referencia, los datos correspondientes a la información de mercado o ratios fundamentales. Existen históricos

de más de veinte años, datos procedentes de búsquedas web o noticias, hay históricos de quince años, procedentes de redes sociales de cinco años, y procedentes de sensores en torno a tres años.

(14) Abreviatura del término robot que se refiere a programas informáticos capaces de efectuar tareas concretas de forma autónoma y análoga a como las realizaría una persona.

(15) Conocido como algoritmo *minimax*.

(16) Aunque el ajedrez tiene un número de jugadas posibles cuantificable, es un número aún demasiado elevado como para poder utilizar estrategias que consideren todo el tiempo todos los casos posibles. En estos casos se emplean metodologías heurísticas.

(17) Este *software* fue desarrollado y liderado por David Silver, Aja Huang y Demis Hassabis.

(18) El término *trading* hace referencia a la operativa de compra y venta de activos financieros. Suele tener una connotación especulativa o cortoplacista, en contraposición con el término inversión.

(19) Aunque las causas del movimiento de mayo de 2016 se investigaron, no se llegó a ninguna conclusión. Tanto es así que se dice que fue provocado por lo que conocemos como *fat finger* en el que, al parecer, un analista de Citygroup se equivocó al pulsar sobre «B» de *billions* (miles de millones) en lugar de pulsar la tecla de «M» de *millions* en una negociación de acciones.

(20) Recordemos el ejemplo de Jeopardy, un algoritmo capaz de interpretar el lenguaje de las preguntas, y con ello, buscar respuestas.

(21) Un ejemplo muy llamativo se produjo en el año 2016, cuando se celebraban las elecciones presidenciales de EE.UU. entre Hillary Clinton y Donald Trump. Las encuestas previas daban ventaja a Hillary, sin embargo, con los primeros recuentos y ante la posibilidad de que Donald Trump ganase, los índices bursátiles cayeron en picado, para luego recuperarse y terminar subiendo, cuando se pasó de posibilidad a realidad.

(22) Estrategia de inversión consistente en comprar activos y mantenerlos en cartera a largo plazo.

(23) Una de las frases más célebres de Warren Buffet establece que el precio es lo que se paga por las cosas y el valor lo que se recibe de ellas. Sin embargo, existen inversores que consideran que no existe diferencia alguna entre precio y valor.

(24) Una operación de arbitraje consiste en la compraventa de un activo en dos mer-

cados distintos, aprovechando la diferencia de precio. Este tipo de operaciones permiten equilibrar los precios de los mercados hasta que desaparece la ineficiencia en la formación del precio.

(25) *Inside trading* hace referencia a la gestión de activos financieros en base a información privilegiada; en muchas legislaciones, es considerada ilegal.

(26) Con el término singularidad tecnológica hace referencia al momento futuro hipotético en el que la inteligencia artificial será autosuficiente y sustituirá al hombre de forma general en la toma de decisiones.

BIBLIOGRAFÍA

DAVEY, K. J. (2014). *Building Winning Algorithmic Trading Systems*. John Wiley & Sons, Inc.

ELDER, A. (1993). *Trading for a Living: Psychology, Trading Tactics, Money Management*. John Wiley & Sons, Inc.

GOODFELLOW, I., BENGIO, Y. y COURVILLE, A. (2016). *Deep learning*. MIT Press.

JOHNSON, B. (2010). *Algorithmic Trading and DMA: An introduction access Trading Strategies*. Barry Johnson.

LECUN, Y., BENGIO, Y. y HINTON, G. (2015). Deep learning. *Nature*, 521(7553), p. 436.

LIU, Y., ZENG, Q., YANG, H. y CARRIO, A. (2018). Stock Price Movement Prediction from Financial News with Deep Learning and Knowledge Graph Embedding. *Knowledge*

Management and Acquisition for Intelligent Systems: 15th Pacific Rim Knowledge Acquisition Workshop, PKAW 2018, Nanjing, China, August 28-29, 2018, Proceedings. Vol. 11016.

MALKIEL, B. G. (2019). *A Random Walk Down Wall Street*. W. W. Norton & Company.

MNIH, V., KAVUKCUOGLU, K., SILVER, D., GRAVES, A., ANTONOGLU, I., WIERSTRA, D. y RIEDMILLER, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

SUTTON, R. S. y BARTO, A. G. (1998). *Introduction to reinforcement learning* (Vol. 2, n.º 4). Cambridge: MIT Press.